VARIABILITY AMONG TEN TALKERS OF WORD INTELLIGIBILITY IN NOISE

by

Barbara Kirk
Susan Marks
and
Russell L. Sergeant

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
NAVAL SUBMARINE MEDICAL CENTER REPORT NO. 693

Bureau of Medicine and Surgery, Navy Department
Research Work Unit M4305.08-3003DAC9.05

Transmitted by:

J Donald Harris

J. Donald Harris, Ph.D.
Head, Auditory Research Branch

Reviewed and Approved by:

Charles F. Gell

Charles F. Gell, M.D., D.Sc.(Med)
Scientific Director
NavSubMedRschLab

Reviewed and Approved by:

Joseph D Bloom.

J. D. Bloom, CDR MC USN
Officer-in-Charge
NavSubMedRschLab

Approved and Released By:

J E Stark

J. E. Stark, CAPT MC USN
COMMANDING OFFICER
NAVSUBMEDCEN

THE PROBLEM

To look at the variance in intelligibility among talkers when noise is mixed with the signal at the listener's ear.

FINDINGS

With individual voices equated for intensity, a range of 28% in words correct was found for normal-hearing listeners. A minimum of five voices was needed properly to sample the variance among talkers. Acoustic and/or perceptual qualities which rendered certain voices more intelligible were discussed but not quantified in this report.

APPLICATION

For communications engineers designing circuit tests using specific word lists and specific talkers.

ADMINISTRATIVE INFORMATION

This investigation was conducted as a part of Bureau of Medicine and Surgery Research Work Unit M4305.08-3003DAC9 and ONR Contract with the University of Connecticut N000014-67-A-0197-0001, both of which are in the area of voice communication by divers and swimmers. The report is number 5 on the work unit and was approved for publication on 12 January 1972. It was designated as NavSubMedRschLab Rpt. No. 693.

Miss Kirk and Miss Marks are graduate students at the University of Connecticut working under the ONR Contract.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL CENTER

ABSTRACT

In an effort to determine what acoustic differences render one voice more intelligible than others, especially in the presence of noise, or reverberation, or when breathing helium gas mixtures, experiments were conducted using ten young adult talkers.

The work of numerous other researchers who have worked with this problem is also reviewed and discussed (32 references).

It is concluded that it may not be good practice to depend upon a single voice in assessing the performance of a communication system. Even when each of ten talkers with dialect-free speech were equated for acoustic intensity, when speaking at a comfortable level, one talker yielded performance of 75.5% words correct, another only 47.5%. It would be necessary to take the average intelligibility of any five of these talkers in order to properly assess a communication system.

More research needs to be done, to study the acoustic and perceptual cues which render a particular voice highly intelligible in the presence of noise.

VARIABILITY AMONG TEN TALKERS OF WORD INTELLIGIBILITY IN NOISE

INTRODUCTION

The speech of some talkers is more intelligible than others. The important questions involved are (1) what are the acoustic differences which render one voice more intelligible than others, and (2) what are the relations among talkers' intelligibility not only in quiet but also in noise, in reverberation, when peak-clipped, when breathing helium, etc. If certain voices could be shown to be especially resistant to all types of distortion, and their invariant acoustic characteristics easily identified, one could select especially suitable (and reject unsuitable) voices by means of a quick acoustic analysis of the voice, and possibly, if the dominant characteristics were open to modification, all or most talkers could be trained to meet certain standards of communication ability.

Variance Among Talkers of Intelligibility in Quiet

In some intelligibility studies conducted by W. B. Snow and A. Meyer at the Bell Telephone Laboratory (see Fletcher and Galt[1]), the variance of level of voice for four men and four women was found to be as great as 16 dB. Certainly these differences contribute profoundly to differences in intelligibility. Fletcher and Steinberg[2] in an even more extensive study also found that the normal talking level of persons varied rather widely, as expected. But when these differences were compensated for by having all talkers use a Volume Indicator while speaking, differences among talkers remained such that a minimum of 5 different normal voices, each using good clear standard general American dialect, were needed to establish mean intelligibility for such voices. They found that when voice level was equated, 21 men were, on the average, a bit more intelligible for nonsense syllables in quiet than 23 women. The classic study of Dunn and White[3] showed that the average speech spectra for 6 men was a bit stronger at 300-3000 Hz than for 5 women. Benson and Hirsh's[4] study reported overall levels about 10 dB lower. Differences between these studies can only be attributed to talker variances.

In their fundamental paper on speech intelligibility, Fletcher and Galt[1] showed that the general talking level (re 10^{-16} watt/cm^2) of 95% of their talkers varied from 55-75 dB around an average of 68 dB. Thus, on speech intensity alone, a wide range of intelligibility among talkers could be expected. Fletcher and Galt also attempted to quantify two acoustic contributions of masking to intelligibility, the residual masking of one speech sound upon another later sound, and the masking of one sound on a simultaneous sound in a different region. Presumably both types of masking will vary among voices with differences in the frequency spectrum and temporal pattern.

Although, just as for isolated vowels, the relationships or differences in frequency among formants is

relatively constant among talkers or for different voice pitches, so also isolated sustained consonants reveal formants which bear certain relatively constant relationships to each other. Note however that in conversation the acoustic structure of the consonant is modified by adjacent speech sounds. In fact, C. M. Harris [5] found that removal of these transitions which characterize a particular voice, rendered it much less intelligible. The problem is therefore acoustically quite complex.

Some attempts have been made to determine just what characterizes an intelligible voice in quiet. Moser et al [6] found that intelligibility of six normal male talkers varied as a function of induced hypernasality or hyponasality, significantly much more so for some talkers than for others. They stated, "Hypernasality might be experienced as a result of foreign language influence or insufficient closure of the nasopharynx; it would be expected as a continuous, albeit perhaps relatively rare, condition associated with a particular speaker. On the other hand, a hyponasal condition caused by the blockage of the nasal passage might occur at frequent intervals for any speaker."

For normal speech in quiet we may conclude that talker differences create ranges of about 20 dB in overall level, but even when this feature is controlled, about 5 voices are needed to sample the remaining differences in intelligibility among good clear talkers.

Variance Among Talkers of Intelligibility in Noise.

During World War II interest was high in predicting and controlling intelligibility not in quiet, but in high-level military noise. It could not be supposed a priori that those qualities which rendered a voice intelligible for nonsense syllables in quiet also allowed it to cut through all noises.

Abrams et al[7] found intelligibility scores of twelve normal-talking young men to range from 44 to 85 per cent correct working against airplane noise. Abrams et al[8] studied the performance of 28 male college students, 10 female talkers (secretaries) with normal speech and 28 untrained enlisted men, against background noise. The sailors yielded standard deviations of about 9 per cent intelligibility with listener panels - with one word list, for example, mean intelligibility was 66%, individual talkers ranging (12 SD) from 47.2 to 84.4%. For 22 men vs 10 women, mean scores were equal (47, 48%), with mean readings on the VU also the same (-3.8, -3.7). Fletcher and Steinberg[2] showed 21 men to be on the average a bit more intelligible than 23 women, for nonsense syllables in quiet.

Abrams et al[9] studied the intelligibility in noise of the speech of 270 college freshman. Their speech professor identified 14 men with the most especially faulty or undesirable speech traits. The following quote from reference 9 shows that these were in fact quite faulty.

"The 14 members of Group P showed no obvious speech faults. Reference to Professor Packard's interview cards for members of Group F showed that they had been selected for the following speech faults:

1. DB -- Sloppy diction; poor clarity; unaccented

syllables; inaudible; lack of confidence; pronounced regional (Michigan) dialect.

2. RC -- Nasility; imperfect enunciation: "These" "Dese", etc., very hesitant; strong racial (French-Canadian) dialect.

3. HD -- Sloppy and careless diction; lack of confidence; high immature nasal voice; lack of force and precision; rate too fast and irregular.

4. GG -- Sloppiest diction immaginable; motionless lips and jaw; bad dental consonants; too rapid rate; no stress; bad regional (South Carolina) dialect.

5. PG -- Sloppy diction; poor clarity; lack of earnestness; slow, monotonous rate.

6. PH -- High pitch; thin vowels; too fast; jumpy inflections; effeminate intonation.

7. SH -- Poor clarity; faulty consonants; L is nasalized; R = W; T not explosive enough; glottal stop used; "these" - "dese", etc.; racial (Jewish) type of phoneme enunication.

8. HH -- Growling tone; very hesitant rhythm; poor L.

9. JK -- Sloppy diction; flat nasal "A"; hesitant; faulty pronunciation ("excape", "goin"); bad reader.

10. RL -- Sloppy, jumbled, elided diction; extreme nasility; hesitant; defective reader.

11. PP -- Nasal; thin vowels; hesitant; faulty pronunciation ("excape", "goin"); bad reader.

12. RS -- Hastily; low depressing tone; indistinct elided diction; faulty pronunciation of stop consonants.

13. ES -- Poor clarity; hesitant; lack of confidence; nasality; regional (Maine) rustic pronunciation; very defective reader.

14. CW -- Very poor clarity and confidence; very hesitant; nasility; meager vocabulary; poor reader."

Nevertheless compared with normal controls, when both groups spoke through low-fidelity circuitry against high level noise, the poor talkers were nearly as intelligible as the good talkers (poor had mn = 56%, S.D. of 8.0%, while good had

mn = 63%, S. D. of 9.8% for words; poor had mn = 66%, S. D. of 9.4 and good had mn = 78%, S. D. of 10.8 for sentence tests). When, to salve the feelings of the Speech Department they were asked to characterize the least intelligible talkers, the speech professors abandoned the terminology shown above for Group F and spoke of "distinctness and force in the utterance of consonants", "maintenance of a deliberate rate of speech", "regularity of loudness-level", "rate of phrasing", and "a clear ringing tone as opposed to a choked muffled, weak tone". An analysis of individual voices showed that poor loudness was a dominant factor, but not exclusively. Of those with poor intelligibility, only one had a loud voice, but this was counterbalanced by poor tone quality and extremely inaccurate consonant articulation. All high-intelligibility talkers had a loud voice, good clear tonal quality, and forceful or exaggerated consonant articulation.

Miller et al[10], correlated intelligibility of 47 talkers against four acoustic indexes of the voice, and against judges' ratings on 9 separate criteria. The most important feature was the speech intensity used, which is not helpful here as presumably the less intelligible talkers could have spoken louder. But by using a multiple-correlation technique, the authors found that a noise-penetrating quality emerged composed of the acoustic cues of intensity, pitch, and "peakedness" and combined with the perceptual cues of judged strength and judged precision of the consonants. Rate in words/min was not important so long as it did not exceed about 120 w/m. On the basis of these data, NSMRL during World War II conducted a telephone talker's school, giving a four-hour course on handling a mike, speaking as loud as possible without strain, and articulating distinctly.

Black and Mason[11] covered substantially the same ground, finding that even when loudness was comparable the intelligibility of a sample of 136 unselected and untrained talkers differed, the standard deviation of differences in intelligibility remaining at about 10% (i.e., intelligibility scores for 2/3 of talkers ranged from 60-80 percent correct). Black[12] has summarized his positive findings on differences in talker intelligibility: (1) intelligibility of talker can be improved by training and/or appropriate feedback; (2) a talker will articulate more clearly in response to a better than to a poorly articulated message.

It is clear that from the data of Abrams and Black speech power is the dominant contributor to intelligibility in noise, but that other features of a voice are also involved. In one study (see figure 5.7 of Hirsh[13]) a recorded list of PB words was displayed on a graphic level recorder, and in a re-recording words which were above the average in power were reduced and those below increased, so that all words were at the same power. Now when the re-recorded words were given to listening panels, rather than homogenizing intelligibility, it turned out that they differed more widely in intelligibility even than before.

Salmon[14] selected words spoken by four talkers with the highest and lowest intelligibility scores from an initial

group of 20. The words were analyzed to determine if the effects of duration, intensity, consonant-vowel ratio (C/V) and peak-clipping on intelligibility. With regard to duration, vocalic sections of the words spoken by the highly intelligible group were not significantly different from the vocalic sections of the group with low intelligibility. Vowels and consonants for the highly intelligible group were on the average 4.0 and 10.3 dB greater than for the low group; however, the group with low intelligibility had averaged C/V ratios of -18.9 against -12.5 dB for the most intelligible group. The results for peak-clipping cannot be generalized; when the C/V ratio exceeded 15 dB, peak-clipping decreased intelligibility. Peak-clipping which resulted in a consonant-vowel difference of less than 3 dB enhanced intelligibility scores.

Goodfriend[15] attempted to measure proficiency of articulatory gesture among talkers, and related articulation to intelligibility. Kelly[16] showed that talkers with longer-duration syllables were more intelligible in noise. Williams et al[17] suggested after studying eight talkers, that intelligibility in noise was related to individual consonant-vowel amplitude ratio. Griffiths et al[18] corroborated this notion for two talkers, and added that the more intelligible talker had more high-frequency energy in consequence of the relatively more powerful consonants in that voice.

Voice Recognition Studies. Communications engineers have often expressed the need for speech analyses which identify the phonemic invariances needed for satisfactory intelligibility, while discarding all that which is redundant. Stevens[19], for example, showed how the short time autocorrelation function can be used in speech analysis, particularly for fricatives. Potter and Steinberg[20] examined with the sound spectrograph and the cathode-ray sound spectroscope the information-bearing aspects of vowels from 76 talkers, including women and children, representing a deliberate attempt to vary pitch, inflection, rate of utterance and vocal cavity dimensions. A large literature has now grown up on the automatic recognition of a particular talker's voice (see Kersta[21] and Pruzansky[22]). It is these advanced techniques, the modern equivalents of the cruder acoustic analysis of Abrams and his colleagues[7-10] which will one day identify those characteristics of a voice which make it especially intelligible in this or that condition of distortion.

An interesting experiment was performed by Peters[23] who determined the degree of each of five types of distortion electronically imposed upon the voice which led to a deterioration of a listener's ability to match a voice to a given sample. Of course, a voice might be quite distinctive (say Louis Armstrong or Mae West) and identifiable under great distortion, but be quite unintelligible under even slight distortion. However, much voice identification research is likely to apply also to intelligibility.

A special problem arises in connection with changes in a voice depending on the emotional state of the talker, on his physiological condition from day to day (state of larynx and nasal cavities, or on the type of noise in the

environment). Of course, speech sounds have acoustic attributes independent of these changes (and even common across all talkers) or the listener could not decode the utterance into linguistic units intended by the speaker; and we are of course not concerned with those acoustic or perceptual qualities of the talker's voice which are free to be actualized differently by a talker on different occasions (recall Chaucer: "Somewhat he lipsed fro his wantonnesse"). We are concerned only with those acoustic perceptual constants which contribute to individual differences in intelligibility in noise. Stevens[24] speculates on the acoustic attributes which appear to be effective in distinguishing one talker from another (and inferentially to intelligibility differences in noise): average fundamental frequency (length and mass of vocal cords); formant frequencies (vocal tract length); spectra of nasal consonants (shape of nasal cavities); spectrum shape of the vowel /1/ in the range of the second, third, and fourth formants (dimensions of oral cavity in relation to length of pharyngeal cavity); and spectrum and intensity of the strident fricative /s/ and /s/ (configuration of hard palate and teeth); plus learned articulatory habits. These are only some of the acoustic features of a voice which may contribute to its intelligibility in noise.

In addition to acoustic analysis of voices, the ear itself can be trusted to assign descriptive labels to a particular voice. From a total of 49 bipolar verbal scales (clear-hazy, rough-smooth, rumbling-shiny, fast-slow, etc.), Voiers[25] winnowed out 4 significant vocal features, using a factor-analytic technique. These features of (1) clarity, (2) roughness, (3) magnitude, and (4) animations were felt by Voiers to exhaust the labels to be put on a voice. It would be important to know how any one or all of these four perceptual attributes is related to intelligibility of a particular voice in noise. If trained listeners could accurately characterize a voice on one or more significant perceptual variables, a tool would be ready at hand to select or reject voices for certain duties or to design and later assess the effect of certain voice-training procedures.

Holmgren[26] showed the relation between certain acoustic measurements of a voice (rate of speaking, mean variance of amplitude of unvoiced sounds, amplitude of voiced sounds, fundamental frequency) and judgments on Voiers' four factors (see also Clarke, Becker and Nixon[27] for further use of Holmgren's perceptual scales). It is clear that at the present time speech science has little notion, even a general notion, of either the acoustic or perceptual qualities of a voice which render it especially intelligible in noise; and such specific questions have hardly even been asked as, whether a voice intelligible in one noise is also relatively intelligible in other noises, or whether a talker can adjust his output to be more intelligible in any particular type of noise.

<u>A Statistical Approach to Variance Among Talkers.</u> By now the instruments and techniques for specifying the acoustic characteristics of an individual voice are in a high state of sophistication, but the contribution to intelligibility of the invariants for intelligibility have not, even as yet, been quantified.

In the meantime, a statistical attack is all that is possible.

Inasmuch as not nearly enough is known about what acoustic cues render a voice more intelligible, so that a "typical" voice cannot be chosen in advance for an intelligibility test, it becomes necessary to select voices in an effort to sample the normal range. Much earlier work on creating intelligibility tests used one trained talker with general American dialect. It is now known that for many purposes this is the worst possible approach. Pollack[28] adopted the expedient of pretesting a number of trained normal talkers in his conditions, then selecting the most and the least intelligible in an attempt to strike a mean. The conclusion of Fletcher and Steinberg[2] has been mentioned that five voices are needed to sample talker variance. No doubt a large number would be needed to represent untrained as well as trained talkers. Harris[29] used an average of four adult male spoken voices, four adult female spoken voices and four female whispered voices in one intelligibility study, and later[30] used five male and five female talkers, all normal but with wide ranges of age and vocal gesture, within a single 50-word predictability test. However, talker differences as such were not reported by either Pollack or Harris.

This paper reports inter-talker variability of intelligibility in noise at S/N = 5 dB for ten young adult talkers, five men and five women.

METHOD

Subjects

Talkers. Five young men and five young women with no obvious speech defects or strong dialects were used.

Listeners. Twenty normal-hearing college students, 10 male and 10 female were used.

Recording the Speech Material

The Modified Rhyme Test (MRT)[31] was selected on grounds of high reliability and ease of administering and scoring.

Two equated 50-word lists of the MRT were taped with a Shure Microphone and an Ampex PR-10 recorder. Each talker enunciated five words on each list, the order of talkers in the second list differing from that in the first. Male and female talkers were interspersed.

Each talker was given a practice period prior to recording. The carrier phrase "Hear the word" prefaced each stimulus word, and a 3-sec pause followed the stimulus word. The talker attempted to maintain a constant level of vocal output by watching a VU-meter. A calibrating tone at VU=0 was placed on each talker's tape.

The tapes from each talker were then played one by one to a General

Radio Graphic Level Recorder, and the average speech power for an individual talker was noted and compared to the mean for all 10 talkers. In a re-recording, each talker's tape was adjusted slightly (using a calibrating tone) so that the average speech power was the same for all talkers. Table 1 shows by how much each talker's tape was adjusted. These adjustments were made to eliminate effects of overall intensity and comfortable talking level as factors which would influence the intelligibility

Table 1. Relative mean graphic level recorder peaks for each talker (datum is average peak for ten monosyllables)

TALKER	SEX	RELATIVE LEVEL IN dB
SM	F	1.6
JR	M	1.0
MK	F	0.8
BK	F	0.7
PL	M	0.2
CS	F	0.2
TM	M	-0.3
EN	M	-0.4
TK	M	-1.5
RG	F	-2.2

scores obtained. White noise was mixed on the final tape at a -5 dB speech-to-noise ratio in order to approximate 60% correct responses by listening panels (see Sergeant and McKay[32] for discussion of S/N ratios and speech intelligibility).

RESULTS AND DISCUSSION

The mean correct scores for the two groups of talkers and the two groups of listeners, by list, are shown in Table 2. The raw scores used in determining these mean values were subjected to an analysis of variance (Bruning and Lintz, 1968) according to talkers, listeners and sex. Sex of the "Listeners" was insignificant as a factor in the intelligibility scores and consequently the data can be collapsed for listeners. The test for list difference was also insignificant and can be similarly collapsed. The interaction between sex of the talker and the list (either A or B) was statistically significant (F of 8.6, df 1/18) and can be attributed to the variation between the larger differences for list B as contrasted with the slight difference observed for list A (Male/Female talker differences were 14% for List A and 4% for List B).

Analysis of variance showed a significant difference between the variances of listeners' responses to the male vs female talkers (F of 47.0, df 1/18). It is logical to assume that this difference in variances was caused by either true talker differences or errors associated with insufficient sampling of the talkers' speech. The second of these was considered most likely for the scores obtained during this study.

Table 2. Talker and listener intelligibility in mean percent correct response, by sex and list. Note: Entries are rounded to nearest whole number.

	Sex	List	LISTENER INTELLIGIBILITY		
			Male	Female	Both Sexes
TALKER INTELLIGIBILITY	Male	A	60	54	57
		B	51	57	54
		A & B	55	56	55
	Female	A	59	62	61
		B	66	70	68
		A & B	62	66	64
	Both Sexes	A	59	58	59
		B	58	63	61
		A & B	49	51	60

Because of the problem in sampling the talkers' speech, a second evaluation was made of the speech material. The talker scores shown in Table 3 are percent correct responses to each talker's words (total responses for each talker was 10 words x 20 listeners, or 200). The range of overall intelligibility, 47.5 - 75.5%, is about what is usually reported in the literature. Difference between means by sex was not significant ($t=1.87$). Apparently limited sampling does not permit sufficiently strong response to the question of difference in the intelligibility of speech of men and women. On the other hand, there is no question that the ranges of intelligibility for the two groups of talkers overlap greatly.

It is concluded that when talkers are limited to a vocabulary of ten different words each, the variance of intelligibility among talkers is not reduced further than commonly reported in the literature, even when a procedure is followed which adjusts each voice for constant overall peak intensity at a comfortable talking level.

It is obvious from the literature, and we see here as well, that many features

Table 3. Intelligibility in percent words correct for individual talkers

Sex	Talker	Intelligibility Score
FEMALE	MK	75.5
	SM	65.5
	BK	65.5
	CS	60.5
	RG	57.5
MEAN		64.9
MALE	TM	67.0
	JR	59.5
	EN	57.0
	TK	50.5
	PL	47.5
MEAN		56.3

of a voice beyond its overall loudness contribute to its being resistant to white noise masking. For example, the voice of MK (Table 3) was about 10 percentage points more intelligible than the next most intelligible voice. Certainly the intensity of this voice is not its most distinctive characteristic. But just which acoustic feature(s) of this voice render it most intelligible, and at the other extreme, which feature(s) render the voice of PL most unintelligible, if not the particular words assigned to these talkers, has not as yet been determined.

A practical question is, how many fewer voices would be needed to sample the intelligibility of these 10 voices. A random selection of triads shows that any 3 voices would yield an average score within the range of the true mean (obtained mean $\pm$ 3 S.E.), but that 5 voices are needed to specify a mean within $\pm$ 1 S.E. of the obtained mean for all 10 talkers. Especially where the same words may be enunciated by all, a maximum of 5 voices should adequately sample the average inteligibility in noise of talkers with clear, unaccented, dialect-free speech.

SUMMARY

Five young men and five young women each tape-recorded 10 words at a comfortable talking level from two lists of the Modified Rhyme Test. The output of each word was examined with a graphic level recorder, and in a re-recording the mean level of each voice was adjusted to the average for all voices. Intelligibility tests in the presence of background noise (S/N was -5 dB) were conducted by loudspeaker simultaneously to 10 young men and 10 young women. There were no differences due to sex of listeners, nor to word list. Mean differences due to sex of talkers were not significant by t-test.

Range of intelligibility among talkers in percent of words correctly perceived was 47.5 - 75.5, even though all voices had been equated for overall intensity. Most of this range could, therefore, not

be attributed to voice intensity or comfortable speaking level, but to some other specific characteristic(s) of the talker's voice. Since no two talkers enunciated the same words, differences in intelligibility residing in the words themselves perturb these data. It was found that no less than any five of these talkers would be required to specify the true mean of all ten voices within one standard error, even though all voices were of clear enunciation, unaccented, and dialect-free.

REFERENCES

1. FLETCHER, H., and GALT, R.H., The Perception of Speech and its Relation to Telephony. J. Acoust. Soc. Am., 1950, 22, 89-151.

2. FLETCHER, H., and STEINBERG, J. C., Articulation Testing Methods. Bell Systems Technical Journal, 1929, 8, 806.

3. DUNN, H. K., and WHITE, S. D., Statistical Measurements on Conversational Speech, J. Acoust. Soc. Am., 1940, 11, 278-288.

4. BENSON, R. W., and HIRSH, I. J., Some Variables in Audio Spectrometry. J. Acoust. Soc. Am., 1953, 25, 499-505.

5. HARRIS, C. M., Building Blocks in Speech, J. Acoust. Soc. Am., 1953, 25, 962-969.

6. MOSER, H. M., DREHER, J. J., and ADLER, Sol, Comparison of Hyponasality, Hypernasality, and Normal Voice Quality on the Intelligibility of Two-Digit Numbers. J. Acoust. Soc. Am., 1955, 27, 872-874 (6 males).

7. ABRAMS, M. H. and KARLIN, J. E., The Problem of Selecting and Training for Communications in Intense Noise. Harvard Psychoacoustic Laboratory, 1942 (PB 19785).

8. ABRAMS, M. H., GOFFARD, S. J., MILLER, J. and STEVENS, S. S., Factors Related to the Intelligibility of Talkers in Noise. Harvard PAL Report IC-60 of 1 February 1944.

9. ABRAMS, M. H., GOFFARD, S. J., MILLER, J., and STEVENS, S. S., Subjective Ratings of the Intelligibility of Talkers in Noise. Harvard Psychoacoustic Laboratory Report IC-67 of 10 April 1944.

10. MILLERS, G. H., ABRAMS, M. H., GOFFARD, S. J., MILLER J., and STEVENS, S. S. Speech in Noise: A Study of the Attributes Determining its Intelligibility. Harvard Psychoacoustic Laboratory, IC-81, 1 September 1944.

11. BLACK, J.W., and MASON, H.M., Training for Voice Communication. J. Acoust. Soc. Am., 1946, 18, 441-445.

12. BLACK, J. W., Intelligibility in Voice Communication. Final Report to USN Special Devices Center, No. 411-1-17 of 22 September 1949.

13. HIRSH, I. J., The Measurement of Hearing, McGraw-Hill, N.Y., 1952.

14. SALMON, R. D., Talker Variation as Related to Intelligibility. Progress Report #31, Commun. Sci. Lab/ONR, Grant N00014-68-0173-0008, 1 December 1970.

15. GOODFRIEND, L. S. Problems in Audio Engineering, Part III, Articulation - its Measurement and its Relation to Intelligibility in the Determination of the Degree of Deafness. Audio. Engng., 1949, 23, 20-21.

16. KELLY, J. C., Syllable Duration and Intensity Related to Intelligibility. USN Special Devices Center Report TR SDC 104-2-15, 1949.

17. WILLIAMS, C. W., WOODS, B., HECKER, M. H., and STEVENS, K. H., Consonant-Vowel Ratio and Speaker Intelligibility, J. Acoust. Soc. Am., 1966, 39, 1257(A).

18. GRIFFITHS, J. D., Rhyming Minimal Contrasts: A Simplified Diagnostic Articulation Test. J. Acoust. Soc. Am., 1967, 42, 236-241.

19. STEVENS, K. N., Autocorrelation Analysis of Speech Sounds. J. Acoust. Soc. Am., 1950, 22, 769-771.

20. POTTER, R. K., and STEINBERG, J. C., Toward the Specification of Speech, J. Acoust. Soc. Am., 1950, 22, 807-820.

21. KERSTA, L., Voice Spectrogram for Unique Personal Identification. Bell Systems Laboratory Rec., 1962, 40, 214-215.

22. PRUZANSKY, S. Pattern-Matching Procedure for Automatic Talker Recognition, J. Acoust. Soc. Am., 1963, 35, 354-358.

23. PETERS, R. W., Studies in Extra-Messages: The Effect of Various Modifications of the Voice Signal Upon the Ability of Listeners to Identify Speakers' Voices. USN School of Aviation Medicine. Report of 1 May 1956.

24. STEVENS, K. N., Sources of Inter-and Intra-Speaker Variability in the Acoustic Properties of Speech Sounds. Program VII International Congress Phonet. Sci., 1971, 81-82 (Abs)

25. VOIERS, W. D., Perceptual Bases of Speaker Identity, J. Acoust. Soc. Am., 1964, 36, 1065-1073.

26. HOLMGREN, G. L., Physical and Psychological Correlates of Speaker Recognition, J. Speech and Hearing Research, 1967, 10, 57-66.

27. CLARKE, F. R., BECKER, R. W., and NIXON, J. C., Characteristics that Determine Speaker Recognition, Stanford Research Institute, Report ESD-TR-66-636 of December 1966, AD-646135.

28. POLLACK, The Effects of High Pass and Low Pass Filtering on the Intelligibility of Speech in Noise, J. Acoust. Soc. Am., 1948, 20, 259-266.

29. HARRIS, J. D., Some Suggestions for Speech Reception Testing. Arch. Otolaryngol., 1949, 50, 388-405.

30. HARRIS, J. D., Combinations of Distortion in Speech: The 25% Safety Factor by Multiple-Cueing. Arch. Otolaryngol., 1960, 72, 227-232.

31. HARRIS, A. S., WILLIAMS, M. et al, Psychoacoustic Speech Tests: A Modified Rhyme Test. Decisions Science Laboratory, Electronic Systems Division of Air Force Command, USAF Report of June 1963.

32. SERGEANT, R. L. and McKay, C., The Intelligibility of Helium-Speech as a Function of Speech-to-Noise Ratio. NSMC Report #555 of 31 October 1968.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Submarine Medical Research Laboratory Naval Submarine Medical Center	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

Variability Among Ten Talkers of Word Intelligibility in Noise

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Interim Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Barbara Kirk
Susan Marks
Russell L. Sergeant

6. REPORT DATE 12 January 1972	7a. TOTAL NO. OF PAGES 13	7b. NO. OF REFS 32
8a. CONTRACT OR GRANT NO. N00014-67-A-0197-0001 b. PROJECT NO. M4305.08-3003DAC9 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) NSMRL Report # 693 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY NAVSUBMEDRSCHLAB & ONR NAVSUBMEDCEN, SUBASE Washington, D.C. Groton, Conn. 06340

13. ABSTRACT

It may not be good practice to depend upon a single voice in assessing the performance of a communication system. Even when each of ten talkers with dialect-free speech, who spoke at a comfortable level, were equated for acoustic intensity, one talker yielded a performance of 75.5% words correct, another only 47.5%. It would be necessary to take the average intelligibility of any five of these talkers properly to assess a communication system. A great deal of research needs to be done in studying the acoustic and perceptual cues which render a particular voice highly intelligible in noise.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Speech Intelligibility Voice Communications Talker Differences Talker Identification Speech in Noise						